ZTF SN Ia DR2: Environmental dependencies of stretch and luminosity of a volume limited sample of 1,000 Type Ia Supernovae

Ginolin, M.[1],[★] ⓘ, Rigault, M.[1] ⓘ, Smith, M.[1],[2] ⓘ, Copin, Y.[1] ⓘ, Ruppin, F.[1] ⓘ, Dimitriadis, G.[3] ⓘ, Goobar, A.[4] ⓘ, Johansson, J.[4] ⓘ, Maguire, K.[3] ⓘ, Nordin, J.[5] ⓘ, Amenouche, M.[6] ⓘ, Aubert, M.[7], Barjou-Delayre, C.[7], Betoule, M.[8] ⓘ, Burgaz, U.[3] ⓘ, Carreres, B.[9],[10] ⓘ, Deckers, M.[3] ⓘ, Dhawan, S.[11] ⓘ, Feinstein, F.[9], Fouchez, D.[9] ⓘ, Galbany, L.[12],[13] ⓘ, Ganot, C.[1], Harvey, L.[3] ⓘ, de Jaeger, T.[8] ⓘ, Kenworthy, W. D.[4] ⓘ, Kim, Y.-L.[2] ⓘ, Kowalski, M.[5],[14] ⓘ, Kuhn, D.[8] ⓘ, Lacroix, L.[4],[8], Müller-Bravo, T. E.[12],[13] ⓘ, Nugent, P.[15],[16] ⓘ, Popovic, B.[1] ⓘ, Racine, B.[9] ⓘ, Rosnet, P.[7] ⓘ, Rosselli, D.[9] ⓘ, Sollerman, J.[17] ⓘ, Terwel, J.H.[3],[18] ⓘ, Townsend, A.[5] ⓘ, Brugger, J.[19], Bellm, E. C.[20] ⓘ, Kasliwal, M. M.[21] ⓘ, Kulkarni, S.[21], Laher, R. R.[22] ⓘ, Masci, F. J.[22] ⓘ, Riddle, R. L.[19], and Sharma, Y.[21] ⓘ

(Affiliations can be found after the references)

ABSTRACT

Context. Type Ia supernova (SN Ia) cosmology will soon be dominated by systematic, rather than statistical, uncertainties, making it crucial to understand the remaining unknown phenomenon that might affect their luminosity, such as astrophysical biases. To be used in cosmology, SN Ia magnitudes need to be standardised, i.e. corrected for their correlation with lightcurve width and colour.

Aims. Here we investigate how the standardisation procedure used to reduce the scatter of SN Ia luminosities is affected by their environment, with the aim to reduce scatter and improve standardisation.

Methods. We first study the SN Ia stretch distribution, as well as its dependence on environment, as characterised by local and global $(g-z)$ colour and stellar mass. We then look at the standardisation parameter α, which accounts for the correlation between residuals and stretch, along with its environment dependence and linearity. We finally compute magnitude offsets between SNe in different astrophysical environments after colour and stretch standardisation, aka steps. This analysis is made possible due to the unprecedented statistics of the volume-limited Zwicky Transient Facility (ZTF) SN Ia DR2 sample.

Results. The stretch distribution exhibits a bimodal behaviour, as previously found in literature. However, we find the distribution to be dependent on environment. Namely, the mean stretch modes decrease with host stellar mass, at a 9.2σ significance. We demonstrate, at the 13.4σ level, that the stretch-magnitude relation is non-linear, challenging the usual linear stretch-residuals relation currently used in cosmological analyses. Fitting for a broken-α model, we indeed find two different slopes between stretch regimes ($x_1 \lessgtr x_1^0$, with $x_1^0 = -0.48 \pm 0.08$): $\alpha_{\text{low}} = 0.271 \pm 0.011$ and $\alpha_{\text{high}} = 0.083 \pm 0.009$, a $\Delta\alpha = -0.188 \pm 0.014$ difference. As the relative proportion of SNe Ia in the high-stretch/low-stretch modes evolves with redshift and environment, this implies that a single-fitted α also evolves with redshift and environment. Concerning the environmental magnitude offset γ, we find it to be greater than 0.12 mag regardless of the considered environmental tracer used (local or global colour and stellar mass), all measured at the $\geq 5\sigma$ level. When accounting for the stretch-non linearity, these steps increase to ~ 0.17 mag, measured with a 0.01 mag precision. Such strong results highlight the importance of using a large volume limited dataset to probe the underlying SN Ia-host correlations.

Key words. Cosmology: dark energy – supernovae: general

1. Introduction

Type Ia supernovae (SNe Ia) are standardisable candles that enabled the discovery of the acceleration of the Universe's expansion in the late 1990s (Riess et al. 1998; Perlmutter et al. 1999). Today, they remain a key cosmological probe, as they can uniquely measure the recent ($z < 0.5$) expansion rate of the Universe and, as such, are central for the derivation of the dark energy equation of state parameter w (Planck Collaboration 2020; Brout et al. 2022), its potential evolution with redshift, and the direct measurement of the Hubble-Lemaître constant H_0 (Freedman 2021; Riess et al. 2022).

The state-of-the-art measurements of cosmological parameters show that a cosmological constant Λ explains the observed properties of dark energy with w compatible with -1 at the 3% precision level (Betoule et al. 2014; Scolnic et al. 2018; Brout et al. 2022). However, the direct measurement of H_0 is incompatible at the 5σ level with the standard model (Λ Cold Dark

Matter, ΛCDM) when the parameters are anchored by early Universe physics (Macaulay et al. 2019; Feeney et al. 2019; Riess et al. 2022). If the latter is not caused by (necessarily multiple, see Riess et al. 2022) sources of systematic uncertainties, this tension would be a sign of new fundamental physics. Yet, no simple theoretical deviation to the fiducial model is able to explain this tension without creating other issues (see Schöneberg et al. 2022, for a recent review). In that context, it is necessary to further investigate the existence of systematic biases that may affect distances derived from SN Ia data.

SNe Ia, as "standardisable" candles, have a natural scatter of ~ 0.40 mag. However, two empirical relations, the so-called slower-brighter and bluer-brighter relations (Phillips 1993; Tripp 1998), make use of SN Ia lightcurve properties to reduce that scatter down to ~ 0.15 mag. The SALT (Guy et al. 2007, 2010; Betoule et al. 2014) lightcurve fitter is the usual algorithm used to estimate SN Ia lightcurve stretch x_1 and colour parameter c (see Kenworthy et al. (2021) and Augarde et al. (in prep) for a recent re-coding).

★ Corresponding author: m.ginolin@ip2i.in2p3.fr

The much reduced scatter provided by these two linear relations makes SNe Ia the best cosmological distance indicator. Yet, only half of this remaining scatter can be explained by known measurement errors or modelling uncertainties. The rest, dubbed "intrinsic scatter", may be due to unknown systematic uncertainties that could bias distances, thus the measurement of cosmological parameters.

It has been demonstrated that SN Ia properties do vary as a function of their environment. The lightcurve stretch, a purely intrinsic property, has been shown to depend on the host environment, such that older and redder environments host on average faster evolving SNe Ia (e.g. Filippenko 1989; Hamuy et al. 1996; Sullivan et al. 2010; Rigault et al. 2020). Since the galactic star formation quickly evolves with redshift (Madau & Dickinson 2014), it has been suggested that SN Ia intrinsic properties evolve with redshift (Howell et al. 2007), as recently demonstrated at the 5σ level by Nicolas et al. (2021). Yet, as long as the stretch brightness dependency is fully captured by the standardisation procedure, such an intrinsic redshift evolution of SN Ia properties should not affect SN Ia cosmology, apart for when the actual stretch distribution is needed, e.g. for selection effect bias corrections (Scolnic & Kessler 2016).

However, it has been shown that SNe Ia from massive hosts are on average brighter after standardisation than these from low-mass hosts (Kelly et al. 2010; Sullivan et al. 2010; Lampeitl et al. 2010; Childress et al. 2013; Rigault et al. 2020). This so called "mass step" is now accounted for in cosmological analyses (Betoule et al. 2014; Scolnic et al. 2018; Brout et al. 2022; Popovic et al. 2024), but its origin is highly debated. Research works suggest that it could be due to progenitor age (e.g. Rigault et al. 2020; Briday et al. 2021; Kim et al. 2018) or dust property variations (e.g. Brout & Scolnic 2021; Popovic et al. 2021, 2023). Understanding the origin of such variations is required for accurate cosmology, since corrections for redshift evolution or sample selection functions may vary with environment.

In this analysis, we investigate the stretch standardisation procedure, its connection with the mass step, and the relation between SN stretch and progenitor age. In a companion paper (Ginolin et al. 2024), we focus on SN Ia colour, its potential connection with dust, and the accuracy of the colour standardisation. Both papers are based on the second data release of the Zwicky Transient Facility (ZTF, Bellm et al. 2019; Graham et al. 2019; Dekany et al. 2020; Masci et al. 2019) Cosmology Science Working Group (ZTF SN Ia DR2, Rigault et al. 2024a; Smith et al. 2025, following the DR1, Dhawan et al. 2022).

This paper starts in Section 2 with a summary of the ZTF SN Ia DR2 release, where we introduce the sample selection used to create a well controlled volume-limited dataset. In Section 3, we study the stretch distribution, where we present a more complex connection between SN stretch and SN environment than what has been previously reported. In Section 4, we investigate in detail the stretch-magnitude relation, that we find to be significantly non-linear. In Section 5, we then investigate magnitude offsets due to SN environment, aka steps, as well as their connection to the non-linearity of the stretch-residuals relation. We test the robustness of our results in Section 6.1, discuss our results in Section 6, and conclude with Section 7.

Except if mentioned otherwise, we use the recent recalibration of the `SALT2.4` lightcurve fitter (Guy et al. 2010; Betoule et al. 2014) from Taylor et al. (2021) as provided by the ZTF SN Ia DR2 release, following Rigault et al. 2024a; Smith et al. 2025.

2. Data

2.1. Zwicky Transient Facility Cosmology DR2

For this analysis, we use the volume-limited ZTF SN Ia DR2 sample presented in Rigault et al. 2024a; Smith et al. 2025.

The initial DR2 sample contains 2663 spectroscopically confirmed SNe Ia passing basic quality cuts: (1) Good lightcurve sampling, i.e., at least seven 5σ flux detections within the -10 to $+40$ days rest-frame phase range, with at least 2 pre-max detections, at least 2 post-max detections, and at least detections in two bands, (2) Stretch $x_1 \in [3, 3]$ measured with a precision better than $\sigma_{x_1} = 1$, (3) Colour $c \in [-0.2, 0.8]$ measured with a precision better than $\sigma_c = 0.1$, (4) Precision on the estimated peak-luminosity time better than $\sigma_{t_0} = 1$ day and (5) `SALT2` lightcurve fit probability greater than 10^{-7}. Finally, to have a volume limited sample, we limit ourselves to SNe Ia having a redshift $z < 0.06$, following the prescription from survey simulations (Amenouche et al. 2024), so that our sample is free from significant non-random selection functions. SNe Ia in the volume-limited sample thus probe the underlying SN Ia population, with no need to model for complex selection function bias correction.

As suggested by Rose et al. (2022), we extend the colour range further than the usual literature cut, from $c < 0.3$ to $c < 0.8$, as we have a significant fraction (10%) of red SNe Ia. In agreement with Rose et al. (2022), we see no behaviour evolution of SNe at $c > 0.3$, and we thus apply a less restrictive cut (see detailed study in Ginolin et al. 2024). In Sect. 6.1, we show that only considering objects with $c < 0.3$ has no significant impact on our results.

We also discard SNe Ia classified by the ZTF SN Ia collaboration as peculiar. As detailed in Dimitriadis et al. (2024), SNe Ia typically classified as peculiar are the 91bg or Ia-CSM subclasses. We however keep the SNe Ia 91t, as they usually pass cosmological cuts. We show in Sect. 6.1 that including the 91bg or discarding the 91t from our sample has no significant impact on our results. We further discard 26 objects with missing host photometry (see Sect. 2.2).

The final volume-limited sample is thus comprised of 945 SNe Ia. Of these, 75% have a redshift coming from host spectral features, mostly from the MOST Hosts Dark Energy Spectroscopic Instrument (DESI) program (Soumagnac et al. 2024), with a typical precision of $\sigma_z \leq 10^{-4}$, while 25% have a redshift derived from SN Ia spectral features. As demonstrated in Smith et al. (2025), these SN Ia-features redshifts are unbiased and have a typical precision of $\sigma_z \leq 3 \times 10^{-3}$.

As mentioned in Rigault et al. (2024a), a non-linearity in the ZTF CCD read-out started to affect the data in November 2019, following an update of the CCD waveforms. This effect, dubbed "pocket effect", is described fully in Lacroix et al. (in prep), and will be corrected for in the upcoming ZTF SN Ia DR2.5. The pocket effect impacts the point spread function, and the amplitude of this effect depends on the signal to noise of a given exposure. The overall effect is of the order of 1% between 15 mag and 19 mag and is colour independent. This is an issue for cosmology, as it prevent us from deriving accurate absolute fluxes, but does not affect the ZTF SN Ia DR2 analysis, as we only use it for self-comparison. However, simulations have shown that, for this DR2, the pocket effect only marginally affects the stretch, while leaving the colour and the Hubble residuals unchanged. For SNe Ia affected by the pocket effect, the stretch x_1 is shifted by $\Delta x_1 = -0.1$, half of the typical x_1 error in the DR2, independently of the true x_1 (cf Rigault et al. 2024a). We show in Sect. 6.3 that our conclusions are not impacted by the pocket ef-

fect, since our results do not significantly vary when splitting our sample between SNe Ia acquired pre- and post-November 2019.

2.2. Local and global host properties

To study the correlation of SN properties with environment, we use the four environmental tracers available in the DR2: stellar mass and colour (`ps1.g-ps1.z` from PanSTARRS Chambers et al. 2016), both local (2 kpc radius around the SN) and global (whole host galaxy). Environmental property estimation is done using the `HostPhot` package (Müller-Bravo & Galbany 2022). It is described in further detail in Smith et al. (2025), along with the parameter distributions. When comparing SNe Ia from these environments, we split them into two subsamples, using the following cuts:

- Global mass: we take the standard literature cut $\log(M_\star/M_\odot)_{\text{cut}}^{\text{global}} = 10$.
- Local mass: we take the median of the local mass distribution $\log(M_\star/M_\odot)_{\text{cut}}^{\text{local}} = 8.9$.
- Local and global colour: we take the gap visible in the bimodal host colour distribution $(g - z)_{\text{cut}} = 1$ mag.

3. Stretch distribution

In this section, we study the distribution of the `SALT2.4` standardisation parameter x_1 (stretch). All fits are done through likelihood minimisation, with the use of `iminuit` (Dembinski & et al. 2020).

3.1. The nearby SN Ia stretch distribution

The ZTF SN Ia DR2 sample stretch (x_1) distribution is shown in the upper panel of Fig. 1. It exhibits a clear bimodal shape, with a low-stretch mode at $x_1 \sim -1.2$ and a high-stretch mode at $x_1 \sim 0.4$, the first mode being approximately twice more populated than the second. As an additional test, we compute differences in Akaike Information Criterion (AIC, Burnham & Anderson (2004)) with other distributions that could match the stretch distribution by eye. The double Gaussian is strongly favoured over a single Gaussian (ΔAIC = 105) and a skewed Gaussian (ΔAIC = 40).

3.1.1. Discussion on SN Ia stretch bimodality

This distribution is very similar to the one from the Nearby Supernova Factory (SNfactory, Aldering et al. 2002; Rigault et al. 2020) dataset studied in detail in Nicolas et al. (2021, hereafter N21). It is also similar to other nearby SN Ia data sets, like the one from Foundation (Foley et al. 2018) and the low-z Pantheon compilation (Scolnic et al. 2018) both studied in Figure 6 of Popovic et al. (2021), or the Supercal compilation (Scolnic et al. 2015) studied in Figure 3 of Wojtak et al. (2023). These latter samples do exhibit a bimodal distribution, but some with both modes equally populated, unlike what we observed with our volume-limited sample. This is likely caused by complex selection effects affecting those low-z samples (see discussion in N21). In contrast, higher redshift samples do not display a clear bimodal distribution. According to N21 and Rigault et al. (2020), this is to be expected. In their model, since low-stretch SNe Ia only originate from old environments, and since the cosmic star formation strongly increases with redshift (Madau & Dickinson

Table 1. Best-fit values for the stretch model presented in Sect. 3.1.1.

Param.	N21 (fiducial)	This work	Difference (σ)
$\overline{x_1^{\text{high}}}$	0.37 ± 0.05	0.42 ± 0.08	0.5
σ_{high}	0.61 ± 0.04	0.54 ± 0.05	1.2
$\overline{x_1^{\text{low}}}$	-1.22 ± 0.16	-1.24 ± 0.18	0.1
σ_{low}	0.56 ± 0.10	0.73 ± 0.09	1.3
r	0.755 ± 0.05	0.59 ± 0.07	1.9

Notes. Means ($\overline{x_1}$) and standard deviations (σ) of the two Gaussian modes of the bimodal stretch model. r is the relative amplitude of each mode.
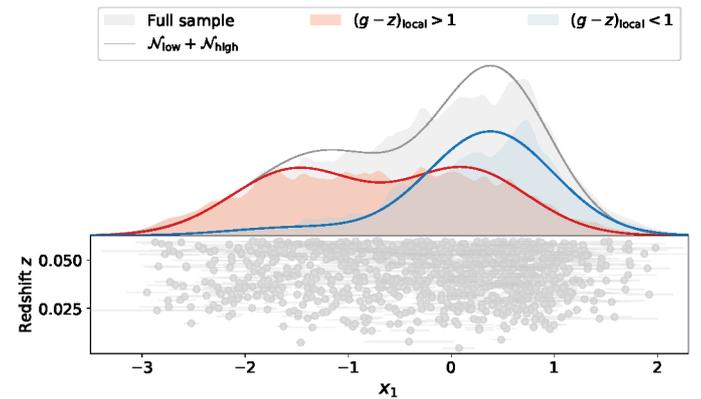
Fig. 1. *Top:* Ideogram of the stretch distribution for the full sample (in grey), and for SNe in locally red/blue environments (509/429 SNe). The full grey line is the bimodal Gaussian from N21 described in Sect. 3, while the red/blue lines are bimodal Gaussian fits to the old/young progenitor subpopulations. *Bottom:* Stretch vs redshift z. The fact that there is no depletion of low stretch SNe at a higher redshift is a consequence of the volume-limited cut.

2014; Tasca et al. 2015), the fraction of (supposedly) old progenitor SNe Ia is lower and, consequently, the low-stretch mode tends to vanish with redshift.

We present in Table 1 the bimodal Gaussian distribution parameters estimated on the volume-limited ZTF data set, and the best fit distribution is shown in Fig. 1. The fitted model is defined as follows: $P(x_1) = r\mathcal{N}(x_1 \mid \overline{x_1^{\text{high}}}, \sigma_{\text{high}}^2) + (1-r)\mathcal{N}(x_1 \mid \overline{x_1^{\text{low}}}, \sigma_{\text{low}}^2)$.

Our results are in good agreements with those reported by N21. To get the ratio of the two modes r for N21, we had to use the assumption made in Rigault et al. (2020) that at their redshift, the fraction of young and old progenitors is half-half. The stretch mode parameters are compatible at the $\sim 1\sigma$ level. The sole remarkable difference is that the amplitude of the r parameter seems lower in our data set, though only at a 1.9σ level. If confirmed, such a reduced amplitude of the high-stretch model could suggest that either the low-stretch mode is slightly more populated than the high-stretch one in the old (delayed) population, or that the fraction of old population is slightly higher than the expected 50% modelled by Rigault et al. (2020) for our median redshift ($z_{\text{median}} = 0.044$).

3.1.2. Stretch distribution in locally red or blue environments

We split the ZTF volume-limited SNe Ia as a function of their local environment to further investigate the origin of the ob-

served bimodality. Following e.g. Roman et al. (2018); Kelsey et al. (2023); Briday et al. (2021); Wiseman et al. (2022), we use the local (2 kpc-radius) colour $(g - z)_{local}$, as it performs well as a proxy for the underlying SN Ia prompt and delayed subpopulations used in N21, with locally blue/red environments hosting young/old SN Ia progenitors. Splitting the ZTF data at $(g - z)_{local} = 1$ mag, 54% of the SNe Ia are found in a locally red environments and 46% in a locally blue environment.

The stretch distribution per local environment and the best bimodal fit for each subgroup is shown in the top panel of Fig. 1. We draw three conclusions from this figure.

First, the locally red environment SNe Ia are consistent with being equally populated with each mode ($r = 49 \pm 9\%$). Second, locally blue environment SNe Ia have a non-null low-stretch mode ($7.5 \pm 1.8\%$). This is consistent with N21 modelling, that assumes the low-stretch mode to only be accessible to old-population SNe Ia once the environmental contamination is accounted for. Indeed, according to Briday et al. (2021), $13^{+7}_{-6}\%$ of the colour classifications are false, i.e. old-progenitor SNe Ia are associated to blue environments and vice versa. We thus expect 8% of old-population SNe Ia to be classified as locally blue (see Fig. 3 of Briday et al. (2021)), and so $\sim 4\%$ of the locally blue sample to be in the low-stretch mode. Third, the mean of the high-stretch mode seems to slightly vary as a function of local environment, with $\overline{\Delta x_1^{high}} = 0.24 \pm 0.15\,(1.6\sigma)$. This is further discussed in the next section.

3.2. Correlation between SN stretch and SN environment

In this subsection, we investigate the correlation of stretch with environment, namely local colour $(g-z)$ and global mass. We are able to disentangle their effects, with the proportion of SNe Ia in high/low stretch modes being dependent on local colour, while the dependence of the mean stretch is tied to global mass. This is then modelled in Sect. 3.2.3.

3.2.1. Qualitative study of stretch environmental correlations

We show in Fig. 2 the connection between SN Ia lightcurve stretch parameter x_1 and host galaxy mass as well as local environmental colour. We clearly see that the low-stretch mode only exists in locally red environments and massive host galaxies. Both are connected, since these environmental parameters are highly correlated, as illustrated in Fig. 3. This is consistent with earlier findings that the low-stretch mode only exists in old stellar population progenitors (e.g. Hamuy et al. 1996; Howell et al. 2007; Rigault et al. 2020; Nicolas et al. 2021; Larison et al. 2024), as those strongly favour massive hosts and make the local environment redder.

Figure 2 further shows that the connection between stretch and environment is more complex than the simple appearance of a low stretch mode in old-population environments as modelled, for instance, by Nicolas et al. (2021). There seems to be a correlation, clearly visible in the high-stretch mode, such that lower mass host/bluer environments tends to have a higher stretch.

To investigate the origin of this effect, we look at the detailed connection between SN lightcurve stretch, global host mass and local environmental colour in Fig. 3.

We split the data along host stellar mass and local environment colour, to investigate which environmental parameters is the most connected the our observations, namely the appearance of the low-stretch mode and the stretch evolution. To do that, we only consider SNe Ia within the 25%-75% range of an envi-
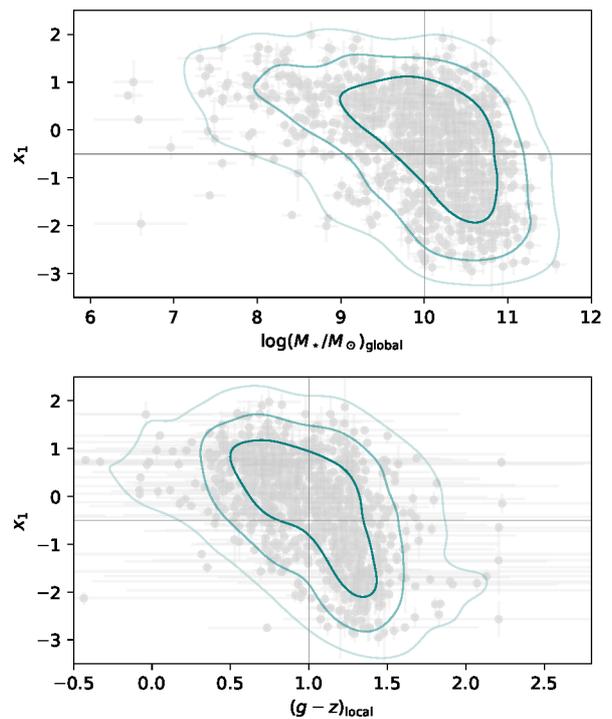
Fig. 2. Correlation between SN Ia lightcurve stretch (x_1) and host stellar mass (*top*), local environmental colour (*bottom*). The contours show the area containing 97%, 84%, and 50% of the SNe Ia. Vertical lines show the environment splitting values, while the horizontal lines show $x_1 = -0.5$, the typical transition point between stretch modes (see Fig. 1).

ronmental parameter (for example host stellar mass), and then study the correlation between x_1 and the other environmental tracer (e.g. local colour), as shown on the top-right panels of Fig. 3. We see that the local colour distribution post host stellar mass 25%-75% selection is similar to that of the entire initial data set. Looking at the bottom-right panel of Fig. 3, we notice the same thing for the stellar mass distribution after the local-environmental 25%-75% selection. However, the resulting x_1 distributions per environment are very informative on the origin of the two observed effect. For the given stellar mass range, locally red SNe Ia seem to have the same high-stretch mode and show in addition the apparition of the low-stretch mode (top-right panel). For a given local environmental colour range, high and low mass host SNe Ia have a similar but shifted ($\Delta x_1 \sim -0.5$) x_1 distributions (bottom-right panel).

From these observations, we conclude that the proportion of SNe in each stretch modes is directly connected to the local colour, while the apparent evolution of the mean stretch is connected to the global host stellar mass. The observed evolution of the high-stretch mode in the bottom panel of Fig. 2 thus seems to be a projection of the evolution connected to the global host stellar mass. Then, assuming that the lightcurve stretch is an intrinsic SN Ia property, we interpret these observations as follows: the low-stretch mode is only accessible to old-population progenitors (delayed) as already suggested in the literature (Howell et al. 2007; Sullivan et al. 2010; Rigault et al. 2020; Nicolas et al. 2021), but the stretch evolution is likely connected to the progenitor metallicity, since stellar mass and stellar metallicity are tightly correlated (Tremonti et al. 2004; Sánchez et al. 2017). This may be connected to literature observations that SN Ia stretch is directly linked to the progenitor mass or pro-

duced ^{56}Ni mass (e.g. Dhawan, Suhail et al. 2017; Scalzo et al. 2014, and references therein).

The link with SN stretch and environment is also studied with the ZTF SN Ia sample using clusters (Ruppin et al. 2024) and voids (Aubert et al. 2024) as environment tracers.

3.2.2. Origin of the shifted stretch distribution between host redshift and SNID redshift SNe Ia

We present in Fig. 4 the SN Ia stretch distribution comparing SNe Ia having a host galaxy redshift ("gal-z") with those without, where the redhsift is extracted from the low resolution SN spectra from the Spectral Energy Distribution machine (SEDm, Blagorodnova et al. 2018), either from SNID (Blondin & Tonry 2007) or from host emission lines ("snid-z", see Smith et al. 2025 for details).

We note that the stretch distribution for "gal-z" SNe Ia is shifted by $\Delta x_1 \sim 0.5$ in comparison to that from "snid-z" SNe Ia. This shift is explained by the selection function caused by our galaxy redshift sources (e.g. the (extended) Baryon Oscillation Spectroscopic Survey ((e)BOSS), see Smith et al. 2025), which strongly favours massive hosts, as illustrated in the top-right panel of Fig. 4. Interestingly, the local colour distributions are similar between these two SN samples. This further supports our claim that the stretch shift as a function of environment is driven by a physical mechanism more directly connected to global stellar mass than local colour, like stellar metallicity (see details in Sect. 3.2.1).

3.2.3. Modelling of the stretch environmental dependencies

To quantify the observations described in Sect. 3.2.1, we extend the stretch model from N21 to account for the global mass dependency of the high/low stretch modes mean $\overline{x_1^{\mathrm{high/low}}}$ described in the previous section, and to get a more robust modelling of the relative fraction of high-stretch SNe Ia r as a function of the local colour. We denote $(g-z)_{\mathrm{local}} \equiv c_{\mathrm{env}}$ and $\log(M_\star/M_\odot)_{\mathrm{global}} \equiv M_{\mathrm{env}}$ for readability. The model is the following:

$$P(x_1 \mid c_{\mathrm{env}}, M_{\mathrm{env}}) = r(c_{\mathrm{env}})\mathcal{N}(x_1|\overline{x_1^{\mathrm{high}}}(M_{\mathrm{env}}), \sigma_{\mathrm{high}}^2)+ \tag{1}$$
$$(1 - r(c_{\mathrm{env}}))\mathcal{N}(x_1|\overline{x_1^{\mathrm{low}}}(M_{\mathrm{env}}), \sigma_{\mathrm{low}}^2)$$

with:

$$r(c_{\mathrm{env}}) = r_{\mathrm{red}} + (r_{\mathrm{blue}} - r_{\mathrm{red}}) \times \mathcal{S}\left(\frac{1}{K_c}[c_{\mathrm{env}} - c_{\mathrm{env}}^0]\right) \tag{2}$$

$$\overline{x_1^{\mathrm{high/low}}}(M_{\mathrm{env}}) = K_M \times (M_{\mathrm{env}} - 10) + \overline{x_1^{\mathrm{high/low,0}}} \tag{3}$$

In Equation 2, $\mathcal{S}$ is a sigmoid function, and r_{red} and r_{blue} correspond to the fraction of high-stretch mode SNe Ia in the blue and red ends of the environmental colour distribution, as illustrated in Fig. 5.

We fit the data with this model accounting for errors on x_1 and c_{env} (i.e., the local colour $(g-z)_{\mathrm{local}}$). The best fit parameters are displayed in Table 2. To quantify the improvement of this model over simpler versions, we compute AIC differences. Comparing N21's model, where the transition between the young and old progenitors stretch distribution (blue and red distribution in Fig. 1) is sharp, with a model with a smooth transition (the sigmoid function plotted in Fig. 5), we find ΔAIC = 29, strongly favouring the smooth transition. We then compare this model to the full model, in which the means of the stretch modes evolves

Table 2. Best fitted parameters for $P(x_1|c_{\mathrm{env}}, M_{\mathrm{env}})$ from Eq. 1-3.

Parameter	Value
K_M	-0.30 ± 0.03
$\overline{x_1^{\mathrm{high,0}}}$	0.23 ± 0.04
σ_{high}	0.54 ± 0.03
$\overline{x_1^{\mathrm{low,0}}}$	-1.34 ± 0.07
σ_{low}	0.64 ± 0.05
r_{red}	0.19 ± 0.07
r_{blue}	0.98 ± 0.03
$(g-z)_{\mathrm{local}}^0$	1.14 ± 0.04
K_c	-0.124 ± 0.028

with mass, we get ΔAIC = 76. The addition of both a continuous dependence of the fraction of SNe Ia in the high-stretch mode on local colour and of the evolution of the means of the stretch modes with global mass is thus justified, as it is strongly supported by the data.

The linear global mass dependency of the stretch modes (K_M) is non-zero at the 9.2σ level, with $K_M = -0.30 \pm 0.03$ dex^{-1}, strongly supporting the evidence that the means of the stretch modes are environment dependent.

We also note that the apparition of the low-stretch mode is progressive, and the fraction of SNe Ia in the low-stretch mode only reaches its redder value of $(1 - r) \sim 0.8$ at $(g - z)_{\mathrm{local}} \sim 1.7$ mag, as visible in Figure 5. This behaviour is more complex than the sharp transition modelled in N21, and explains the $\sim 50\%$ of low-stretch SNe Ia in red environments seen in Fig. 1.

4. Stretch-residuals relation

In this section, we study the SN Ia stretch standardisation procedure accounting for the brighter-slower Phillips (1993) relation. We briefly introduce SN Ia standardisation in Sect. 4.1, and present our fitting procedure in Sect. 4.2, as well as the resulting standardisation parameters in Sect. 4.3. We then assess the universality of this procedure as a function of SN environment in Sect. 4.4, to finally question its assumed linearity in Sect. 4.5.

The colour standardisation is studied in detail in a companion paper (Ginolin et al. 2024).

4.1. SNe Ia standardisation

We define the difference between observed and modelled distance moduli, also called Hubble residuals, as:

$$\Delta\mu = \mu_{\mathrm{obs}} - \mu_{\mathrm{cosmo}}, \tag{4}$$

where μ_{cosmo} is calculated in `astropy` (Astropy Collaboration et al. 2013, 2018), following a flat ΛCDM cosmology given by Planck Collaboration (2020, $\Omega_m = 0.315$), plus a blinded magnitude offset. We then use Chauvenet's criterion to reject outliers of the cosmological fit, which discards 7 SNe.

The usual standardisation formula for SNe Ia is given by:

$$\mu_{\mathrm{obs}} = m_B - M_0 - \beta c + \alpha x_1 - \gamma p + \Delta_b \tag{5}$$
$$= \mu - (\beta c - \alpha x_1 + \gamma p - \Delta_b)$$

where M_0 is the absolute B-band SN magnitude (degenerate with H_0), α and β are the linear standardisation coefficients correcting for the stretch (x_1) and colour (c) SN Ia variations following the slower-brighter and bluer-brighter relations (Tripp 1998).
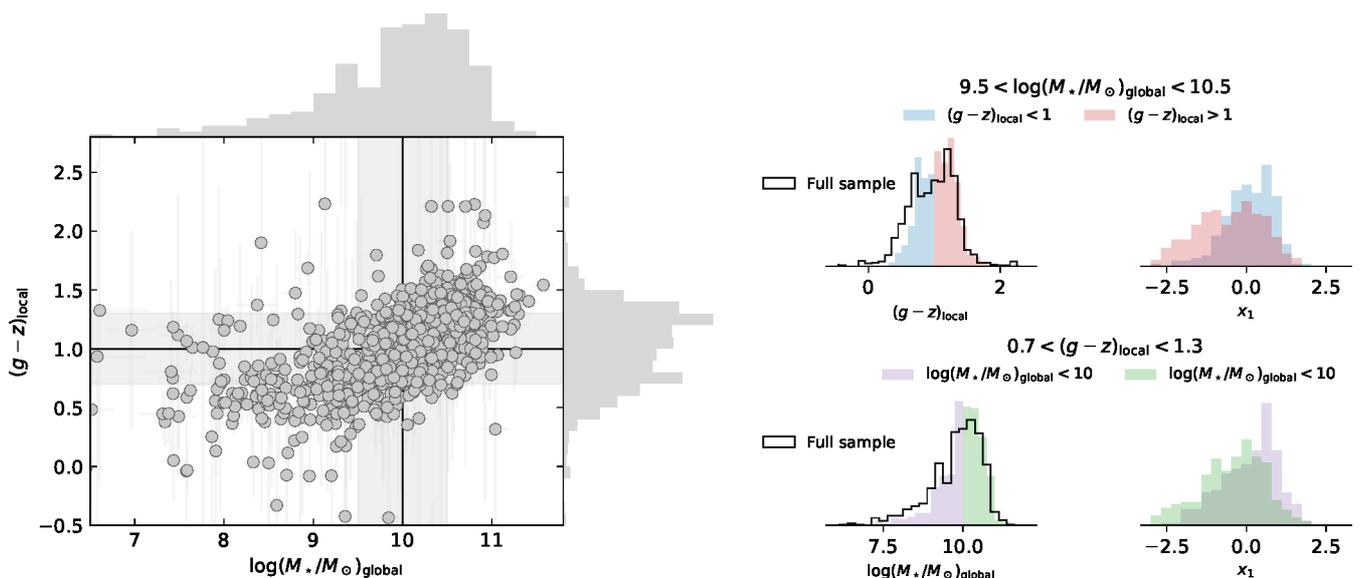
Fig. 3. Connections between SN Ia lightcurve stretch (x_1), global stellar mass ($\log(M_*/M_\odot)_{\text{global}}$) and local environmental colour ($(g-z)_{\text{local}}$), illustrating the complex correlation between stretch and SN environmental properties. *Left*: Local colour vs. global host stellar mass. The light grey band shows the 25%-75% percentile range for host global mass (vertical, [9.5, 10.5] dex) and local environmental colour (horizontal, [0.7, 1.3] mag). These cuts are used to select SNe Ia shown in the right panels. *Top-right:* Local environmental colour and SN stretch distributions for SNe Ia whose host global mass is in [9.5, 10.5] dex. Blue and red histograms show locally blue and red environment SNe Ia. The black histogram is the local colour distribution for the full sample. *Bottom-right:* Global host stellar mass and SN stretch distributions for SNe Ia whose local environmental colour is in [0.7, 1.3] mag. Purple and green histograms show low and high mass hosts SNe Ia.
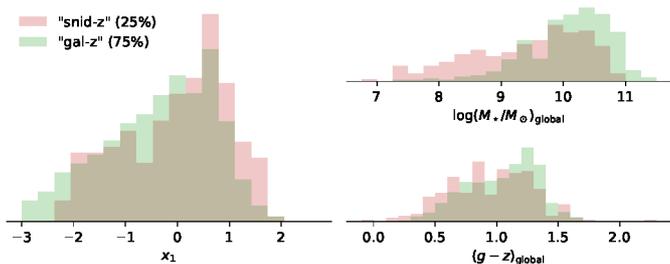
Fig. 4. SN Ia lightcurve stretch (x_1), global host mass ($\log(M_*/M_\odot)_{\text{global}}$) and local environmental colour ($(g-z)_{\text{local}}$) distributions split per redshift origin. SNe Ia with galaxy redshifts are plotted in green ($\sigma z \sim 10^{-4}$), and SNe Ia relying on low-resolution SN spectra (extracted from SEDm, Blagorodnova et al. (2018)), either from SNID (Blondin & Tonry 2007) or on host emission lines are plotted in red ($\sigma z \sim 10^{-3}$). The apparent stretch offset is explained by the selection function associated with galaxy redshifts, which favours high-mass hosts.
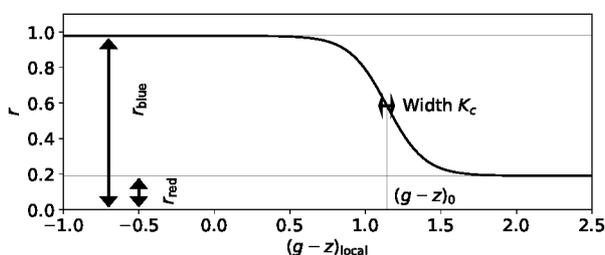
Fig. 5. Fraction of SNe in the high stretch mode as a function of the local colour, as modelled in Eq. 2.

The γp term accounts for SN Ia magnitude environmental dependencies (e.g. Kelly et al. 2010; Sullivan et al. 2010; Briday et al. 2021). p is the probability that an environmental tracer m

is below a given splitting value ("cut", see Table. 3), while γ is the magnitude offset between the SN Ia subpopulations below or above the environment cut, aka the "step". In practice, p ($\in [0, 1]$) is the cumulative distribution function of the environmental proxy m measured with an error δm evaluated at the cut value, such that $p = \int_{-\infty}^{\text{cut}} \mathcal{N}(x, m, \delta m)dx$. In recent cosmological analyses, m is usually the global host stellar mass and γ is referred to as the mass step.

The Δ_b term accounts for selection function affecting the survey, since overly bright objects are easier to acquire and to classify. The correct modelling procedure of this term is highly discussed, and has been shown to bias the environmental correction if not accurately done (e.g. Smith et al. 2020; Popovic et al. 2021; Nicolas et al. 2021; Wiseman et al. 2022). In this analysis, to avoid such complications, we use the volume-limited sample of the ZTF DR2 sample ($z < 0.06$), which is free from non-random selection functions either from lightcurve estimation or spectral typing (Smith et al. 2025; Amenouche et al. 2024). Consequently, we set $\Delta_b = 0$.

The standardisation (i.e. estimation of M_0, α, β, γ, cf. Eq. 5) is done using total-χ^2 minimisation. The total-χ^2 approach enables to fit a model explaining an observed y-variable (here $\Delta\mu$) that depends on input noisy x-variables (here c, and x_1, p). To do so, and unlike simple χ^2 minimisation, one has to fit for the true (noisy) x-variables that are used, in turn, to estimate y.

4.2. Total-χ^2 minimisation.

In practice, for a sample containing N SNe, we fit for $3 \times N + 5$ parameters: the $3 \times N$ parameters corresponding to the true values of the observed noisy standardisation variables (c^{true}, x_1^{true} and p^{true}), the four standardisation parameters of interest from Eq. 5 (α, β, γ and M_0), plus an intrinsic magnitude scatter σ_{int} to account for leftover dispersion in the residuals.

The "total" χ^2 is thus the sum of two parts, one quantifying the likelihood that the observed x correspond to the fitted x_{true}:

$$\chi^2_{\text{param}} = \sum_i \frac{(x_1^{\text{true,i}} - x_1^i)^2}{\sigma_{x_1}^i{}^2} + \sum_i \frac{(c^{\text{true,i}} - c^i)^2}{\sigma_c^i{}^2} + \sum_i \frac{(p^{\text{true,i}} - p^i)^2}{\sigma_p^i{}^2}$$

(6)

and the usual χ^2 standardisation, computed with the "true" x-variables:

$$\chi^2_{\text{res}} = \sum_i \frac{(\mu^i - \beta c^{\text{true,i}} + \alpha x_1^{\text{true,i}} + \gamma p^{\text{true,i}})^2}{\sigma_{\text{obs}}^i{}^2 + \sigma_{\text{int}}^2}$$

(7)

where $\sigma_{\text{obs}}^i{}^2 = \begin{pmatrix} 1 & \alpha & \beta \end{pmatrix} \mathbb{C}^i \begin{pmatrix} 1 \\ \alpha \\ \beta \end{pmatrix}$, and $\mathbb{C}^i$ is the covariance matrix between (μ^i, x_1^i, c^i). Hence, including the determinant $\log\det = \sum_i \log(\sigma_{\text{obs}}^i{}^2 + \sigma_{\text{int}}^2)$, we minimize $\chi^2 = \chi^2_{\text{res}} + \chi^2_{\text{param}} + \log\det$.

In our fit, we consider p as non-noisy as it already takes into account errors on the environment parameter chosen to compute the magnitude step, so we assign it arbitrarily small errors σ_p, such that p^{true} is effectively equal to p. To avoid biases, the intrinsic scatter is fitted iteratively. We fix a given σ_{int}, find the parameters $(\alpha, \beta, \gamma, M_0, (x_1^{\text{true}})_i, (c^{\text{true}})_i)$ that minimise the total-χ^2, then fix those parameters and fit σ_{int} as the value that normalises the reduced total-χ^2. This procedure is repeated 10 times. The error estimation on the parameters is then done at fixed σ_{int} using a 1,000 steps MCMC. Finally, the corrected residuals are computed using the measured c, x_1 and p.

Based on realistic simulations, we show in Appendix A that this fitting approach is unbiased, unlike the simple χ^2 approach, which assumes $x_{\text{obs}} = x_{\text{true}}$. For a wide range of α, β, γ and σ_{int} combinations, we recover the input parameters and accurately estimate their errors, as our residual pulls (difference between the fitted and input parameter divided by the fitted error) are Gaussianly distributed, with a distribution centred on zero with a width of 1.

This fitting procedure makes use of `jax` (Bradbury et al. 2018). Additional details on the use of total-χ^2 for cosmological parameter inference in SN cosmology will be given by Kuhn et al. (in prep).

4.3. Standardisation parameter coefficients

Fitting on our 938 SNe Ia sample with the total-χ^2 procedure described in Sect. 4.2, we find $\alpha = 0.161 \pm 0.010, \beta = 3.05 \pm 0.06$ and $\gamma = 0.143 \pm 0.025$ using local colour as our environmental tracer (e.g., Roman et al. 2018; Kelsey et al. 2021; Briday et al. 2021). σ_{int} is not released, as it will be the subject of a future detailed analysis.

The correlation matrix of the fitted $(\beta, \alpha, M_0, \gamma)$ is plotted in Fig. 6. We see a strong correlation between the stretch standardisation parameter α and the environmental step γ. This is expected, as γ traces the magnitude bias due to environment and stretch is strongly linked with environment, as seen in Sect. 3. Because of the correlation between α, β and γ, it is necessary to fit them jointly. Indeed, fitting for only α and β and computing γ as an a posteriori difference between the Hubble residuals for SNe in different environments will lead to α and β absorbing part of the magnitude-environment connection. This will in turn bias
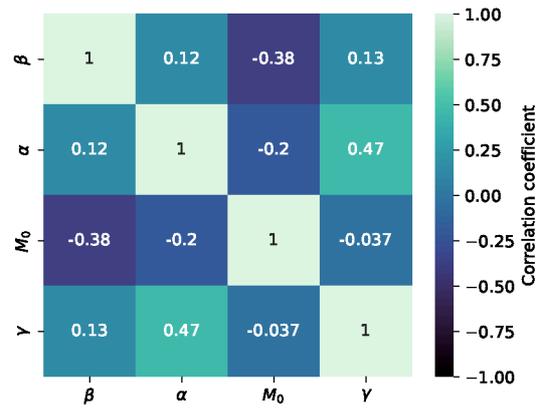

Fig. 6. Correlation matrix of the standardisation parameters from Eq. 5.
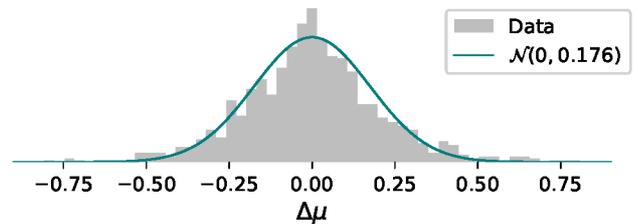

Fig. 7. Histogram of the standardised Hubble residuals (corrected for SN colour, SN stretch and environment). The blue line is a Gaussian centred on 0 with a standard deviation corresponding to the normalised median absolute deviation $\sigma_{\text{nmad}} = 0.176$.

α and β, and underestimate the step γ. This is discussed in more details in Sect. 5.

As illustrated in Fig. 7, the Hubble residuals are normally scattered, with a width (normalised median absolute deviation, nmad) of 0.175 mag. The regular STD of the Hubble residuals is 0.226 mag. Discarding SNe Ia that have redshift extracted from SN Ia spectroscopic features (25% of the sample, hence leaving 707/945 SNe Ia) leads to very similar results ($\alpha = 0.173 \pm 0.009$, $\beta = 3.06 \pm 0.05$ and $\gamma = 0.143 \pm 0.022$), but with a reduced nmad of 0.16 mag (0.213 for the regular STD). This reduction is expected since SN Ia-features redshift having a typical precision of 3×10^{-3} (see details in Smith et al. 2025), corresponding to an additional 0.08 mag scatter, to be added in quadrature.

The amplitude of the step γ is discussed in Sect. 5, β is studied in a companion paper (Ginolin et al. 2024) and α is further discussed in the following subsections.

4.4. Environmental dependency of the stretch standardisation

The universality of the brighter-slower and brighter-bluer empirical linear relations might be challenged by observed environmental dependencies of such standardised SN Ia magnitudes, e.g. the mass-step (e.g. Sullivan et al. 2010), the local colour bias (e.g. Roman et al. 2018) or the age bias (Rigault et al. 2020), that most likely are different aspects of the same underlying effect (e.g. Briday et al. 2021; Brout & Scolnic 2021).

In this subsection, we thus test if the stretch standardisation coefficient α itself is environment dependent. Using the environmental tracers introduced in Sect. 2.2, we split our volume-limited sample in two, following the cuts introduced in that section (see also Table 3). For each of these four environmental

Table 3. Stretch correction coefficient (α) and absolute magnitude (M_0) as a function of the SN environment.

Tracer	Cut	$\Delta \alpha$	ΔM_0
$(g-z)_{\mathrm{local}}$	1 mag	$-0.032 \pm 0.021\ (1.5\sigma)$	$0.123 \pm 0.021\ (5.9\sigma)$
$(g-z)_{\mathrm{global}}$	1 mag	$-0.022 \pm 0.020\ (1.1\sigma)$	$0.112 \pm 0.022\ (5.0\sigma)$
$\log(M_\star/M_\odot)_{\mathrm{local}}$	8.9	$-0.041 \pm 0.019\ (2.2\sigma)$	$0.128 \pm 0.022\ (5.9\sigma)$
$\log(M_\star/M_\odot)_{\mathrm{global}}$	10	$-0.059 \pm 0.020\ (3.0\sigma)$	$0.135 \pm 0.021\ (6.3\sigma)$

tracers, we independently standardise the two SN subsamples. We report in Table 3 the recovered stretch (α) and absolute magnitude (M_0) parameters. The colour term (β) is discussed in a dedicated paper (Ginolin et al. 2024). The M_0 offset, equivalent to the step, is discussed in Sect. 5.

We see in Table 3 that the stretch standardisation parameter α does not seem to strongly depend on the environment. The most significant difference appears when using global host mass as tracer, with $\Delta \alpha \sim -0.06$ at the 3.0σ level, such that high-mass galaxies host SNe Ia with a larger α coefficient. This is consistent with earlier findings that the standardisation coefficients do not seem to depend on environment given the small data sets these analyses had access to (e.g. Sullivan et al. 2010; Rigault et al. 2020; Kelsey et al. 2021). The small dependence of α on environment is actually the signature of the non-linearity of the stretch-residuals relation, studied in detail in Sects. 4.5 and 6.4.

4.5. Linearity of the stretch-magnitude relation

The next standardisation assumption we challenge is the linearity of stretch-residuals relation (see e.g. Figure 8 of Sullivan et al. 2011; Wang et al. 2006; Rubin et al. 2015). More recently Garnavich et al. (2023); Larison et al. (2024) presented 3 to 4.5σ evidence for different α coefficients as a function of SN stretch, using $O(100)$ SNe Ia.

We illustrate in Fig. 8 the stretch standardisation step for our sample. This figure shows the Hubble residuals standardised for colour and environment, but not for stretch, as a function of stretch (x_1). This relation was assumed to be linear by all SN cosmological analyses so far (e.g. Riess et al. 1998; Perlmutter et al. 1999; Betoule et al. 2014; Scolnic et al. 2018; Brout et al. 2022). Yet, it is visible in Fig. 8 that the stretch-residuals relation is not linear, as suggested by Garnavich et al. (2023); Larison et al. (2024). It indeed has a distinct broken shape, with lower-stretch SNe Ia having a stronger (in absolute value) α coefficient than high-stretch SNe Ia. Hints of this behaviour are also seen in the ZTF SN Ia DR2 siblings (Dhawan et al. 2024) and host-morphology (Senzel et al. 2024) studies.

To quantify this observation, we update Eq. 4.1 to allow for a broken-α standardisation. This introduces two extra parameters α_{low} and α_{high}, in place of a single α, and the breaking point x_1^0, which is also fitted:

$$\mu_{\mathrm{obs}} = m_B - M_0 - \beta c + \mathcal{A}(x_1)x_1 + p\gamma \tag{8}$$

with

$$\mathcal{A}(x_1) = \begin{cases} \alpha_{\mathrm{low}} & \text{if } x_1 < x_1^0 \\ \alpha_{\mathrm{high}} & \text{otherwise} \end{cases} \tag{9}$$

Fitting this 'broken line' model, we find $\alpha_{\mathrm{low}} = 0.271 \pm 0.011$ and $\alpha_{\mathrm{high}} = 0.083 \pm 0.009$. The two α significantly differ at the 13.4σ level, with $\Delta \alpha = \alpha_{\mathrm{high}} - \alpha_{\mathrm{low}} = -0.188 \pm 0.014$. Such a

high-significance demonstrates that indeed, the stretch-residuals relation is non-linear. We also find a reduction of the fitted intrinsic scatter (-40%). The reduction of the scatter indicates that the broken-α absorbs some of the unexplained scatter dubbed as intrinsic. While fitting for a broken-α, we find $\beta_{\mathrm{broken}} = 3.31 \pm 0.03$. A detailed study of the intrinsic scatter will be the subject of a dedicated analysis and the actual σ_{int} values are thus not published.

The evaluation of the improvement of the fit by adding two extra parameters is not straightforward, as the χ^2 of the fit depends on σ_{int}, as explained in Sect. 4.1, which differs in the linear and broken-α case. If we fix σ_{int} to the broken-α value, we find ΔAIC = 101. When assuming the larger σ_{int} associated to the linear-α model, we find ΔAIC = 48, with the significance of the broken-α being reduced to 6.2σ. In both cases, the broken-α model is favoured by the data.

We highlight that such a broken-α result is very unlikely to be caused by an unknown selection function. Indeed, this would cause to miss objects preferentially in the top-left corner of Fig. 8 (faster and fainter objects), which would in turn pull α_{high} to lower values. The robustness of this result is further demonstrated in Sect. 6.1.

The best-fit value of the stretch split x_1^0 is also of interest. When fitting the broken-α model with a variable breaking point, we find $x_1^0 = -0.48 \pm 0.08$, which corresponds to the transition between the two stretch modes visible in the x_1 distribution in Fig. 1. This may suggest that there is a physical difference between the two stretch modes, resulting in the different stretch-residuals correlation. With the current modelling used, both modes share the same absolute standardised magnitude at the breaking point, hinting at a continuous transition between the two modes. This would need to be further investigated in a dedicated modelling analysis. Additionally, we investigate in Sect. 6.2 the dependence of the broken-α on SN colour.

4.6. Environmental dependency of the stretch-magnitude relation non-linearity

In Fig. 8, we plot binned environment-standardised SN Hubble residuals as a function of stretch for locally red environment SNe Ia only ($(g-z)_{\mathrm{local}} > 1$). Unlike the blue (younger) environments, these SNe Ia populate both stretch modes (see Sect. 3), which enables us to test if indeed the broken relation is driven by the SN Ia stretch mode (it should thus also be there for red-environment SNe Ia) or if it is an artifact of an environmentally dependent α (the magnitude-stretch relation should then be linear). Considering only red-environment SN Ia, we find $\Delta \alpha = -0.255 \pm 0.048\ (5.3\sigma)$ demonstrating that the broken-α effect is not caused by an environment-dependent α, but is rather a stretch-mode driven effect, with each stretch modes having their own stretch-magnitude relation. We reach the same conclusion when fitting a broken-α on SNe Ia in high-mass hosts, with $\Delta \alpha = -0.204 \pm 0.043\ (4.8\sigma)$. Since the relative fraction of each mode is strongly environment-dependent (cf. Sect. 3), this non-linearity leads to variation of the usual α standardisation as a function of the environment and redshift (see dedicated discussion in Sect. 6.4).

5. SN Ia standardised magnitudes (steps)

In this section, we investigate the amplitude of the environmental magnitude offsets affecting our volume-limited sample. These offsets between SN environments, aka steps, are represented by the γ term in Eq. 5 and computed simultaneously with
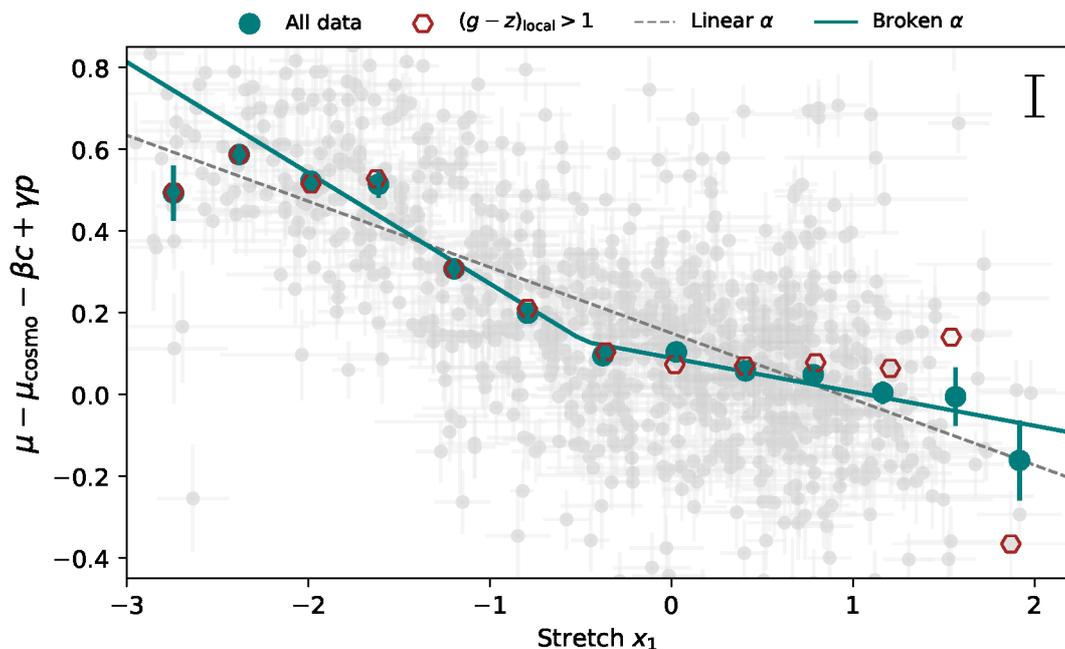
Fig. 8. Standardised Hubble residuals ($\alpha = 0$) as a function of stretch (x_1). Large blue points show the data binned by stretch (mean) while the red hexagons show the binned stretch for SNe Ia from locally red environments. The errorbars on the binned points only display the error on the mean, but a reference for the fitted intrinsic scatter $\sigma_{\rm int}$ is plotted in the top right corner. The blue line shows the best fitted broken-α model, while the dashed grey line shows the best α when assuming the linear Tripp relation.

the other standardisation parameters. We illustrate in Fig. 9 the local colour step, manually setting $\gamma = 0$ to get Hubble residuals uncorrected for environment. This figure clearly demonstrates the existence of astrophysical biases affecting stretch and colour standardised SN Ia magnitudes. Indeed, SNe Ia in locally blue environments are significantly fainter ($\gamma = 0.143 \pm 0.025$ mag) than those from locally red environments. Accounting for the non-linearity of the stretch-residuals relation, we find $\gamma = 0.175 \pm 0.010$ mag.

Nonetheless, the actual origin of the observed astrophysical bias is still highly debated, and Brout & Scolnic (2021); Popovic et al. (2021) suggest that the mass step originates from different dusts properties for different environments (see further discussion in Kelsey et al. 2023; Wiseman et al. 2022). Modelling dust with BayeSN (Thorp et al. 2021; Mandel et al. 2022), Grayling et al. (2024) find an intrinsic mass step of -0.049 ± 0.016 mag. Finally, Smith et al. (2020) shows that selection function corrections inaccuracy affects the ability to measure environmental biases, since all parameters are correlated (e.g. stretch, colour, host mass, local colour). The strength of our volume-limited data set is that we are free from such selection issues.

The fact that the steps, derived from a low-redshift volume-limited dataset, are found to be significantly larger than those from higher-redshift samples may be due to two things. First, the amplitude of the mass step may decrease with redshift (Rigault et al. 2013, 2020; Childress et al. 2014). Second, as the effect of selection function and the dependencies of SNe Ia with their astrophysical environment are intertwined, bias corrections, which are necessary at high-redshift, might affect the derivation of the step. We note that the amplitude of our step is compatible with the value found by Rigault et al. (2020) using another low-redshift, volume-limited sample (SNfactory, Aldering et al. 2002).

Table 4. Amplitude of the environment magnitude offset, aka steps, for the joint fit for (α, β, γ), without and with broken-α.

| Tracer	γ [mag]	$\gamma_{\rm broken-\alpha}$ [mag]
$(g-z)_{\rm local}$	0.143 ± 0.025 ($5.6\,\sigma$)	0.175 ± 0.010 ($18.4\,\sigma$)
$(g-z)_{\rm global}$	0.120 ± 0.022 ($5.5\,\sigma$)	0.149 ± 0.009 ($16.0\,\sigma$)
$\log(M_\star/M_\odot)_{\rm local}$	0.120 ± 0.022 ($5.5\,\sigma$)	0.149 ± 0.010 ($14.9\,\sigma$)
$\log(M_\star/M_\odot)_{\rm global}$	0.145 ± 0.021 ($6.8\,\sigma$)	0.153 ± 0.013 ($12.0\,\sigma$)

The amplitude of the step is connected to the ability to accurately perform stretch and colour standardisation, since SN stretch (mostly) and colour are correlated with environmental parameters (see Fig. 2 for the stretch-local colour connection). Hence, as already discussed in Smith et al. 2020; Dixon 2021; Briday et al. 2021, when fitting first for α and β, and then for γ (as e.g. in Kim et al. 2019 for a recent example), the colour and stretch standardisation will absorb part of the (ignored) astrophysical biases, thus biasing all terms (α, β and γ ; see, e.g. discussion in Rigault et al. 2020; Smith et al. 2020; Dixon 2021; Briday et al. 2021). Such "a posteriori" measurements can then only underestimate the step.

The best fitted γ parameters are presented in Fig. 10 and reported in Table 4. For comparison, the figure also includes the "a posteriori" step results, as well as the ΔM_0 offset acquired when applying a per-environment standardisation and the steps obtained using a broken-α relation (see Sect. 4.5).

All of magnitude offsets in Table 4 are significantly non-zero at the $\geq 5\sigma$ level. When using the usual Tripp (1998) linear standardisation formalism, the strongest one is the global mass step with $\gamma = 0.145 \pm 0.021$ mag ($6.8\,\sigma$). The local colour step is at a similar level, in agreement with Roman et al. (2018), Kelsey et al. (2021), and Briday et al. (2021). The amplitude of the
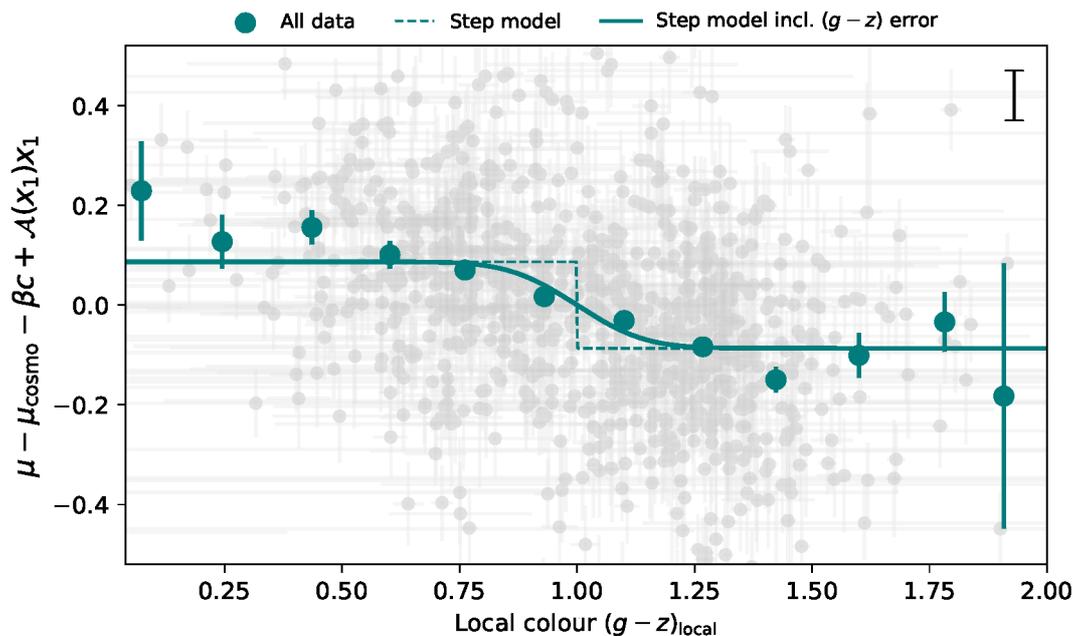
Fig. 9. Standardised Hubble residuals corrected for stretch (with a broken-α) and colour as a function of local environmental colour. Large blue points show the residuals binned by local colour, with the errorbars corresponding to the error on the mean. The scale of the fitted intrinsic scatter is shown in the top right corner. The dashed line shows the step model. The full line illustrates the step model once convolved by a Gaussian with a standard dispersion equals to the mean local environmental measurement error ($\sigma_{(g-z),\mathrm{local}} \sim 0.14$), as we account for those errors in the step model.
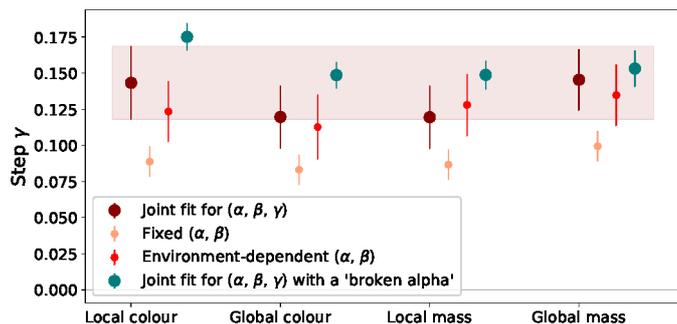
Fig. 10. Fitted steps for each of the environmental proxies. The dark red and green points indicate steps where both (α, β) and γ are fitted at the same time for each proxy. The light red points indicate the steps when the same (α, β) correction (fixed across proxies) is applied to the whole sample. The red points indicate the steps when independent (α, β) corrections are applied for each of the subsamples. The green points are steps when including the broken-α model into the fit, according to Sect. 4.5. The shaded band represents the 1σ interval for the reference step (local colour step from the joint fit).

global mass step is consistent with findings reported in SNfactory (Rigault et al. 2020), but higher than the steps reported in SDSS/SNLS (Roman et al. 2018) and DESY3 (Kelsey et al. 2021).

As expected, Fig. 10 shows that the environmental steps derived after standardisation are significantly biased towards smaller values, with a reduction of ~ 0.05 mag). When comparing the amplitude of magnitude offsets when the standardisation is made independently for the two environments, all steps remain strongly significant ($> 5\sigma$ level), demonstrating that environmental inhomogeneity of the standardisation procedure is not responsible for SN magnitude biases, as already suggested

in Sect. 4.4. Furthermore, accounting for the non-linearity of the stretch-magnitude relation (cf. Sect. 4.5) slightly increases the amplitude of all the environmental steps and reduce the parameter errors, see Fig. 10. This is to be expected if indeed the non-linearity is true, since the fitted model is more representative of the data, and thus all parameters are better measured. This also demonstrates that the existence of two stretch modes having each their own α term is not responsible for the environmental magnitude offsets. Altogether, local tracers perform slightly better than global ones when using the broken-α standardisation.

To ensure that the discrepancies between the step values are indeed due to variations in the fitting procedure, we investigated the differences between these step fitting methods using simulations in Sect. A.3. We reproduce both the order of the points and the factor between the jointly-fitted step (red points) and the a-posteriori step (light red points) for the local colour step. This confirms that not performing a joint fit for the environmental step along with α and β will underestimate it.

6. Discussion

6.1. Robustness tests

In this section, we test the robustness of our main results, namely the evolution of the stretch modes with global mass (K_M parameter from Eq. 3), the non-linearity of the stretch-residuals relation and the amplitude of the magnitude steps, by applying several analysis variations:

– Varying the redshift cut we use to define our volume limited sample, reducing it to $z_{\max} = 0.05$ as a conservative cut to test if our results could be caused by leftover selection function. We also extend our analyses to $z_{\max} = 0.07$, to gain in statistical power, with a limited Malmquist bias.

**Table 5.** Robustness tests as described in Sect. 6.1.

Test	N_{SN}	K_M [σ]	$\Delta\alpha$ [σ]	γ [mag]	$\gamma_{broken-\alpha}$ [mag]
Fiducial	945	9.2	13.4	0.143 ± 0.025	0.175 ± 0.010
$z_{max} = 0.05$	613	7.9	11.9	0.164 ± 0.028	0.204 ± 0.011
$z_{max} = 0.07$	1341	10.6	14.7	0.143 ± 0.017	0.182 ± 0.007
Host z	707	6.9	9.7	0.143 ± 0.024	0.163 ± 0.011
incl. Ia-pec	965	9.7	13.7	0.145 ± 0.022	0.182 ± 0.010
no 91t	878	8.8	15.0	0.150 ± 0.023	0.172 ± 0.012
$c < 0.3$	852	9.1	10.0	0.144 ± 0.022	0.167 ± 0.012
$\chi^2_{SALT} < 0.1$	735	8.4	12.9	0.148 ± 0.021	0.182 ± 0.011
Before Oct. 2019	451	5.4	7.3	0.132 ± 0.032	0.170 ± 0.011
After Dec. 2019	420	6.8	11.8	0.162 ± 0.033	0.201 ± 0.013

Notes. Number of SNe in the sample, significance of the evolution of the stretch modes with colour (K_M parameter from Eq. 3), difference in α when fitting a broken standardisation law (using a variable cut) and step in local $(g - z)$ (with and without broken-α standardisation), for each of the tests performed in Sects. 6.1-6.3.

– Only using SNe Ia with host redshifts, whose residual scatter is smaller.

– Changing the SN Ia subpopulations used in the study. We include the subluminous 91bg population as well as SNe Ia classified as peculiar, as these could be hard to identify in higher redshift surveys for instance, or in surveys relying on photo-typing. We also redo the analysis discarding 91t. This allows us to test that our results are not driven by the 91t subpopulation.

– Using the literature cut on SN colour ($c < 0.3$), to make sure our results are not affected by the inclusion of red objects.

– Applying a much stronger cut on lightcurve fit probability ($\chi^2_{SALT} < 0.1$), to check if our results could be induced by ill-modelled lightcurves.

The result of all these variations are summarised in Table 5.

We see no significant variations across all modifications made. Accessing an even more secured volume limited sample only slightly reduces the significance of our results, as we reduce our statistics, strengthening our claim that the broken-α or the step amplitudes are not caused by any uncontrolled selection effect leftover. Going up to $z_{max} = 0.07$ does not changes our results, which might hint that the volume-limited cut we made was conservative. Only using host redshifts reduces the significance of the evolution of stretch with global mass. This is expected, as the host redshift sample is biased towards higher mass hosts (cf. Sect. 3), thus reducing the leverage of the fit. The significance of the broken-α also decreases, due to the reduced statistics. Adding the peculiar 91bg-like objects, discarding 91t from our sample or only using $c < 0.3$ SNe does not impact the significance of our results. The significance of the broken-α decreases slightly when discarding $c > 0.3$ objects, but is still higher than the 10σ level. This confirms that red SNe do not exhibit peculiar behaviour, confirming the conclusions of Rose et al. (2022) and Ginolin et al. (2024). Enforcing a strong lightcurve fit requirement also has only a slight impact on K_M and $\Delta\alpha$, due to lower statistics, confirming that it is not a lightcurve modelling issue that may cause either the broken stretch-residuals relation or the observed astrophysical bias (see a dedicated lightcurve residual analysis in Rigault et al. (2024b).
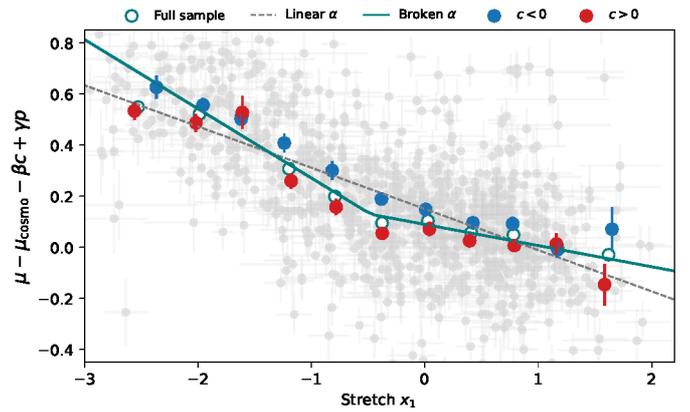
Fig. 11. Hubble residuals corrected for colour and environment as a function of stretch x_1 (same as Fig. 8), for the full sample (white points) and blue/red SNe (blue/red points).

6.2. Dependence of the broken-α on SN colour

As an additional test of the broken-α, we looked at its value when cutting the sample into red ($c > 0$, 559 SNe) and blue ($c < 0$, 379 SNe) SNe. We show in Fig. 11 the residuals corrected for colour and environment against stretch. The broken-α is still visible for both subsamples. We find $\Delta\alpha$ to be respectively $8.5\sigma/5.1\sigma$ away from 0 for the red/blue subsample. The reduction of the significance is expected due to the lower number of SNe in each subsample. We thus conclude that the non-linearity of the stretch-residuals relation does not depend on SN colour.

We note in Fig. 11 that the blue SNe Ia residuals are ~ 0.07 mag fainter than the red SNe Ia ones. As detailed in Appendix A of Ginolin et al. (2024), this offset is caused by measurement noise in colour, that creates an up-tail towards blue values on correctly standardised Hubble residuals. We highlight that this is not a signature of a non-linearity in β; see Sect. 4.1 of Ginolin et al. (2024), which confirms that the colour-residuals relation is linear.

6.3. Pocket effect

As discussed in Sect. 2, the "pocket effect" that appeared in some of the ZTF CCDs after October 2019 can impact stretch measurements (Lacroix et al. (in prep), Rigault et al. 2024a), leading to an overall shift in stretch of $\Delta x_1 \sim -0.1$ for these SNe Ia.

In this section, we investigate the impact of this effect on our results by cutting our sample in two: SNe with t_0 before the 1st of October 2019 (451 SNe Ia), thus unaffected by the pocket effect, and those with t_0 after December 2019 (420 SNe Ia). SNe Ia at peak in October and November are discarded, as their lightcurves are only partially affected.

We first check that our data is indeed impacted by a stretch shift. We thus refit the double Gaussian stretch model presented in Sect. 3.1, but allowing for a shift between the two samples. The recovered parameters of the double Gaussian are similar to those from the full sample (within 1σ) and the fitted shift is $\Delta x_1 = -0.1 \pm 0.06$, in agreement with simulations.

To check that our main results are unaffected by the pocket effect, we then report in Table 5 the values of the evolution of the stretch modes with host mass, the significance of the broken-α and the environmental steps for SNe Ia before/after the pocket effect, as was done in Sect. 6.1.

We first look at the dependency of the stretch mode means with the host stellar mass, parametrised by K_M. The significance of K_M reduces from 9σ (full sample) to $\sim 6\sigma$ for SNe Ia from both subsamples. Such a lower significance is expected given the reduction of the sample size. Indeed, randomly drawing 450 SNe Ia from the full sample and refitting for K_M leads to a 6σ or lower signal 36% of the time. Moreover, the values of K_M are similar between subsamples ($K_M = -0.33 \pm 0.05$ for SNe Ia after Dec. 2019 and $K_M = -0.24 \pm 0.04$ for SNe Ia before Oct. 2019), and similar to the value for the full sample ($K_M = -0.30 \pm 0.03$).

The amplitude of the broken-α shows the strongest variation between the two samples: 7.3σ for the 451 SNe Ia unaffected by the pocket effect and 11.8σ for the 420 SNe Ia impacted by it. It thus raises the question if this difference is simple causes for a reduction of the sample size. To evaluate this, we randomly select 450 SNe Ia from the full sample and fit for a broken-α. Doing this process 1,000 times, we find 22% of the downsized sample have $\Delta\alpha < 7.3\sigma$ (96% of 11.8σ of less). Furthermore, the values of $\Delta\alpha$ are compatible between subsamples: for the SNe Ia before Oct. 2019, we find $\alpha_{\text{low}} = 0.257 \pm 0.014$ and $\alpha_{\text{high}} = 0.131 \pm 0.010$, while for the SNe Ia after Dec. 2019 we find $\alpha_{\text{low}} = 0.300 \pm 0.014$ and $\alpha_{\text{high}} = 0.086 \pm 0.011$. We thus conclude that the pocket effect does not significantly impact the broken-α.

We finally investigate the environmental step amplitudes. The steps are similar between samples (less than a $\sim 1.5\sigma$ difference), both for the linear and broken-α standardisation. We thus conclude that the pocket effect has no significant impact on the amplitude of the steps reported in this paper.

6.4. Broken α and linear standardisation biases

As presented in Sect. 4.5, the Phillips (1993) stretch-residuals relation, parametrised by the α term in Eq. 5, is non-linear. This non-linearity coincides with the lightcurve stretch bimodality (cf. Fig. 1 and, e.g. N21), such that high-stretch mode SNe Ia have a weaker stretch-residuals correlation ($\alpha_{\text{high}} \sim 0.08$) than those from the low-stretch mode ($\alpha_{\text{low}} \sim 0.27$, see Fig. 1 and dedicated discussion in Sect. 4.5). Yet, the low-stretch mode is only accessible to SNe Ia from old population environments (see Sect. 3.1.2 and e.g. Howell et al. 2007, Sullivan et al. 2010, Rigault et al. 2020, Nicolas et al. 2021), unlike the high-stretch mode, which is always available. Since the fraction of old-population SNe Ia varies with redshift and/or with environment, the use of a (usual) linear standardisation parameter α will lead to variations of α with redshift and environment. This is particularly important in two cases, redshift evolution and selection bias.

We here evaluate qualitatively the effect of fitting a linear α if the truth is a broken α. A more thorough investigation of this problem, along with simulations, will be presented in a future paper.

Following Rigault et al. (2020), the fraction of SNe Ia from young progenitors increases with redshift, going to almost 1 at $z \sim 2$. As they only populate the high-stretch mode, which has $\alpha_{\text{high}} \sim 0.08$, a linear α will evolve from its main sample value ($\alpha \sim 0.16$) to a value close to α_{high} when going to higher redshifts. We thus predict a decrease of a linear α with redshift.

Following the same reasoning, as low-stretch SN Ia are fainter, lowering the magnitude limit of a survey thus increases the fraction of SNe Ia in the high-stretch mode, and biases a linear α towards α_{high}. We thus predict a decrease of a linear α when discarding objects due to a magnitude limit. Additionally, as shown in Sect. 3, stretch is strongly correlated with environment. Indeed, SNe Ia in locally blue environments and/or low

mass hosts only have access to the high-stretch mode. Hence, selecting SNe Ia only in those environments would bias a linear α towards α_{high} (and towards α_{low} for locally red environments and/or high mass hosts). Yet, it is common for SN Ia samples to be biased towards a given environment, e.g. mid-mass blue spiral galaxies, as it is easier to get a host redshift, massive galaxies, as it is easier to get a host mass, or targeted surveys that only observe massive spirals, to maximise the acquired SNe Ia rate. The effect of magnitude and environment are likely to add up, making disentangling their contributions even harder. This thus calls for caution when fitting a linear stretch-residuals standardisation, especially on non volume-limited surveys. A more qualitative study of these effects, as well as the impact on the resulting fitted cosmology, will be presented in a upcoming analysis.

7. Conclusion

This paper presents a detailed analysis of the SN Ia standardisation procedure, focusing on the stretch (x_1) residuals relation and SN environment magnitude offsets, aka steps. The companion paper Ginolin et al. (2024) focuses on SN colour.

This analysis is based on the volume-limited data set ($z < 0.06$) of ~ 1000 SN Ia from the SN Ia ZTF DR2 sample (Rigault et al. 2024a; Smith et al. 2025). This data set is free from significant selection function and therefore observed distributions, correlations and biases are representative of the true underlying SN Ia nature.

Our conclusions are the following:

1. The stretch-residuals relation, represented by the α term in the SN Ia standardisation (Tripp 1998), is non-linear. Low-stretch SNe Ia ($x_1 \leq -0.48$) have a much stronger magnitude-stretch relation with $\alpha_{\text{low}} = 0.271 \pm 0.011$ than high-stretch mode SNe Ia $\alpha_{\text{high}} = 0.083 \pm 0.009$. This non-linearity is measured at the 13.4σ level.

2. The environmental dependency of standardised SN Ia magnitude is stronger than usually claimed. Accounting for the stretch-residuals non-linearity we find $\gamma = 0.175 \pm 0.010$ mag using the local (2 kpc) colour ($g - z$) tracer, a 18.4σ detection. Even with a usual linear α (Tripp 1998) relation, we find $\gamma = 0.143 \pm 0.025$ mag (a 5.6σ detection).

3. The global mass step is at a similar level than the local environmental colour step, with $\gamma = 0.145 \pm 0.021$ mag (0.153 ± 0.013 mag, using a broken-α).

4. The stretch distribution is bimodal with mode parameters measured in close agreement with those derived by Nicolas et al. (2021). The low-stretch mode ($x_1 \leq -0.5$) is only available to SNe Ia from old-population environments (redder galaxies and/or more massive host), unlike the high-stretch mode.

5. The stretch modes means decrease linearly as a function of global host stellar mass (at a 9.2σ significance). We show that this, in turn, causes an apparent evolution of the stretch mode with local colour. This change is likely caused by the progenitor metallicity, that more directly correlates with stellar mass than with local environmental colour.

6. The complex stretch-environment connection can be summarised as follows: SNe Ia have two stretch modes, each having their own α. The low stretch mode is only available to SNe Ia from old-population environments (i.e. to "delayed" SNe Ia). The mean of the x_1 distribution slowly evolves as a function of host stellar mass, suggesting that SN metallicity influences x_1, such that higher metallicity (more massive host) leads to SNe Ia with smaller stretch.

7. Because the stretch distribution evolves with redshift and as a function of environment, we expect that the Tripp (1998) linear α term should decrease as a function of redshift. This should also strongly vary as a function of selection effect, such as magnitude limit or those favouring for instance blue emission line galaxies and/or mass hosts.

This analysis of the ZTF volume-limited sample has revealed unexpected SN Ia behaviours, notably the strong non-linearity of the SN Ia standardisation procedure. This would likely not have been possible without the combination of a large sample statistic and the fact that our sample is not significantly affected by any (potentially hard to model) selection function effects. As such, we are able to predict what such effects could cause on the derivation of the SN Ia standardisation parameters that, in turn, leads to the inference of cosmological parameters. Our well controlled sample also reveals large environmental SN magnitude offsets. Altogether, those astrophysical biases, induced by our limited understanding of what SNe Ia truly are, are a dominating source of systematic uncertainties, along with photometric calibration. To mitigate them, we strongly recommend for future surveys like the Legacy Survey of Space and Time (LSST) and the Nancy Grace Roman Space Telescope to dedicate a fraction of their observing time to build a volume-limited SN Ia data set covering as much of the SN parameter phase space as possible. This is particularly true for the redshift component as redshift dependent issues directly translate into biases on the derivation of cosmological parameters.

References

Aldering, G., Adam, G., Antilogus, P., et al. 2002, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4836, Survey and Other Telescope Technologies and Discoveries, ed. J. A. Tyson & S. Wolff, 61–72

Amenouche, M., Smith, M., Rosnet, P., et al. 2024, A&A Submitted (ZTF SI)

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Aubert, M., Rosnet, P., Popovic, B., et al. 2024, A&A Submitted (ZTF SI)

Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002

Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22

Blagorodnova, N., Neill, J. D., Walters, R., et al. 2018, PASP, 130, 035003

Blondin, S. & Tonry, J. L. 2007, ApJ, 666, 1024

Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, JAX: composable transformations of Python+NumPy programs

Briday, M., Rigault, M., Graziani, R., et al. 2021, A&A, 657, A22

Brout, D. & Scolnic, D. 2021, ApJ, 909, 26

Brout, D., Scolnic, D., Popovic, B., et al. 2022, ApJ, 938, 110

Burnham, K. P. & Anderson, D. R. 2004, Sociological Methods & Research, 33, 261

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560

Childress, M., Aldering, G., Antilogus, P., et al. 2013, ApJ, 770, 108

Childress, M. J., Wolf, C., & Zahid, H. J. 2014, MNRAS, 445, 1898

Dekany, R., Smith, R. M., Riddle, R., et al. 2020, PASP, 132, 038001

Dembinski, H. & et al., P. O. 2020

Dhawan, S., Goobar, A., Smith, M., et al. 2022, MNRAS, 510, 2228

Dhawan, S., Mortsell, E., Johansson, J., et al. 2024, A&A Submitted (ZTF SI)

Dhawan, Suhail, Leibundgut, B., Spyromilio, J., & Blondin, S. 2017, A&A, 602, A118

Dimitriadis, G., Burgaz, U., Deckers, M., et al. 2024, A&A Submitted (ZTF SI)

Dixon, S. 2021, PASP, 133, 054501

Feeney, S. M., Peiris, H. V., Williamson, A. R., et al. 2019, Phys. Rev. Lett., 122, 061105

Filippenko, A. V. 1989, PASP, 101, 588

Foley, R. J., Scolnic, D., Rest, A., et al. 2018, MNRAS, 475, 193

Freedman, W. L. 2021, ApJ, 919, 16

Garnavich, P., Wood, C. M., Milne, P., et al. 2023, ApJ, 953, 35

Ginolin, M., Rigault, M., Copin, Y., et al. 2024, A&A Submitted (ZTF SI)

Graham, M. J., Kulkarni, S. R., Bellm, E. C., et al. 2019, PASP, 131, 078001

Grayling, M., Thorp, S., Mandel, K. S., et al. 2024, MNRAS, 531, 953

Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11

Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7

Hamuy, M., Phillips, M. M., Suntzeff, N. B., et al. 1996, AJ, 112, 2438

Howell, D. A., Sullivan, M., Conley, A., & Carlberg, R. 2007, ApJ, 667, L37

Kelly, B. C. 2007, ApJ, 665, 1489

Kelly, P. L., Hicken, M., Burke, D. L., Mandel, K. S., & Kirshner, R. P. 2010, ApJ, 715, 743

Kelsey, L., Sullivan, M., Smith, M., et al. 2021, MNRAS, 501, 4861

Kelsey, L., Sullivan, M., Wiseman, P., et al. 2023, MNRAS, 519, 3046

Kenworthy, W. D., Jones, D. O., Dai, M., et al. 2021, ApJ, 923, 265

Kim, Y.-L., Kang, Y., & Lee, Y.-W. 2019, Journal of Korean Astronomical Society, 52, 181

Kim, Y.-L., Smith, M., Sullivan, M., & Lee, Y.-W. 2018, ApJ, 854, 24

Kowalski, M., Rubin, D., Aldering, G., et al. 2008, ApJ, 686, 749

Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, ApJ, 722, 566

Larison, C., Jha, S. W., Kwok, L. A., & Camacho-Neves, Y. 2024, ApJ, 961, 185

Macaulay, E., Nichol, R. C., Bacon, D., et al. 2019, MNRAS, 486, 2184

Madau, P. & Dickinson, M. 2014, ARA&A, 52, 415

Mandel, K. S., Thorp, S., Narayan, G., Friedman, A. S., & Avelino, A. 2022, MNRAS, 510, 3939

Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, PASP, 131, 018003

Müller-Bravo, T. & Galbany, L. 2022, The Journal of Open Source Software, 7, 4508

Nicolas, N., Rigault, M., Copin, Y., et al. 2021, A&A, 649, A74

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Phillips, M. M. 1993, ApJ, 413, L105

Planck Collaboration. 2020, A&A, 641, A6

Popovic, B., Brout, D., Kessler, R., & Scolnic, D. 2023, ApJ, 945, 84

Popovic, B., Brout, D., Kessler, R., Scolnic, D., & Lu, L. 2021, ApJ, 913, 49

Popovic, B., Scolnic, D., Vincenzi, M., et al. 2024, MNRAS, 529, 2100

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Riess, A. G., Yuan, W., Macri, L. M., et al. 2022, ApJ, 934, L7
Rigault, M., Brinnel, V., Aldering, G., et al. 2020, A&A, 644, A176
Rigault, M., Copin, Y., Aldering, G., et al. 2013, A&A, 560, A66
Rigault, M., Smith, M., Goobar, A., et al. 2024a, A&A Submitted (ZTF SI)
Rigault, M., Smith, M., Regnault, N., et al. 2024b, A&A Submitted (ZTF SI)
Roman, M., Hardin, D., Betoule, M., et al. 2018, A&A, 615, A68
Rose, B. M., Popovic, B., Scolnic, D., & Brout, D. 2022, MNRAS, 516, 4822
Rubin, D., Aldering, G., Barbary, K., et al. 2015, ApJ, 813, 137
Ruppin, F., Rigault, M., Ginolin, M., et al. 2024, A&A Submitted (ZTF SI)
Sánchez, S. F., Barrera-Ballesteros, J. K., Sánchez-Menguiano, L., et al. 2017, MNRAS, 469, 2121
Scalzo, R., Aldering, G., Antilogus, P., et al. 2014, MNRAS, 440, 1498
Schöneberg, N., Abellán, G. F., Sánchez, A. P., et al. 2022, Phys. Rep., 984, 1
Scolnic, D., Casertano, S., Riess, A., et al. 2015, ApJ, 815, 117
Scolnic, D. & Kessler, R. 2016, ApJ, 822, L35
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Senzel, R., Maguire, K., Burgaz, U., et al. 2024, A&A Submitted (ZTF SI)
Smith, M., Rigault, M., Dimitriadis, G., et al. 2025, A&A Submitted (ZTF SI)
Smith, M., Sullivan, M., Wiseman, P., et al. 2020, MNRAS, 494, 4426
Soumagnac, M. T., Nugent, P., Knop, R. A., et al. 2024, ApJS, 275, 22
Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782
Sullivan, M., Guy, J., Conley, A., et al. 2011, ApJ, 737, 102
Tasca, L. A. M., Le Fèvre, O., Hathi, N. P., et al. 2015, A&A, 581, A54
Taylor, G., Lidman, C., Tucker, B. E., et al. 2021, MNRAS, 504, 4111
Thorp, S., Mandel, K. S., Jones, D. O., Ward, S. M., & Narayan, G. 2021, MNRAS, 508, 4310
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Tripp, R. 1998, A&A, 331, 815
Wang, L., Strovink, M., Conley, A., et al. 2006, ApJ, 641, 50
Wiseman, P., Vincenzi, M., Sullivan, M., et al. 2022, MNRAS, 515, 4587
Wojtak, R., Hjorth, J., & Hjortlund, J. O. 2023, MNRAS, 525, 5187

[1] Univ Lyon, Univ Claude Bernard Lyon 1, CNRS, IP2I Lyon/IN2P3, UMR 5822, F-69622, Villeurbanne, France
[2] Department of Physics, Lancaster University, Lancs LA1 4YB, UK
[3] School of Physics, Trinity College Dublin, College Green, Dublin 2, Ireland
[4] Oskar Klein Centre, Department of Physics, Stockholm University, SE-10691 Stockholm, Sweden
[5] Institut für Physik, Humboldt Universität zu Berlin, Newtonstr 15, 12101 Berlin, Germany
[6] National Research Council of Canada, Herzberg Astronomy & Astrophysics Research Centre, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada
[7] Université Clermont Auvergne, CNRS/IN2P3, LPCA, F-63000 Clermont-Ferrand, France
[8] Sorbonne Université, CNRS/IN2P3, LPNHE, F-75005, Paris, France
[9] Aix Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
[10] Department of Physics, Duke University, Durham, NC 27708, USA
[11] Institute of Astronomy and Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[12] Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193, Barcelona, Spain
[13] Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain
[14] Deutsches Elektronen Synchrotron DESY, Platanenallee 6, 15738, Zeuthen, Germany
[15] Lawrence Berkeley National Laboratory, 1 Cyclotron Road MS 50B-4206, Berkeley, CA, 94720, USA
[16] Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA
[17] Oskar Klein Centre, Department of Astronomy, Stockholm University, SE-10691 Stockholm, Sweden
[18] Nordic Optical Telescope, Rambla José Ana Fernández Pérez 7, ES-38711 Breña Baja, Spain
[19] Caltech Optical Observatories, California Institute of Technology, Pasadena, CA 91125, USA
[20] DIRAC Institute, Department of Astronomy, University of Washington, 3910 15th Avenue NE, Seattle, WA 98195, USA
[21] Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
[22] IPAC, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA

Appendix A: Fitting procedure

As mentioned in Sect. 4.1, we use a "total-χ^2 approach to fit for standardisation. This is motivated by the known bias introduced by the usual χ^2 method when fitting for a line where both axis have measurement errors. This issue is described in Kelly (2007), and detailed in Kowalski et al. (2008) in a cosmological case (see also Appendix A of Ginolin et al 2024b for an example case on SNe standardisation).

In this section we present realistic simulations of our dataset on which we perform both the regular log-likelihood minimisation (simple χ^2) and the total χ^2 method used in this paper. These simulation are made using skysuvey[1]. We compare in Fig. A.1 our data with skysurvey simulations. In this analysis, to be able to quickly simulate variants of the survey, we bypass the lightcurve fitting step. We tested that for ZTF-like survey this method does not introduce any significant biases on the SN parameters or their errors. Both the errors and the correlations between lightcurves parameters and environment proxies are indeed realistic. A detailed study of the underlying SNe Ia population using such realistic simulations is ongoing and will be the subject of a dedicated follow-up analysis.

Appendix A.1: Total χ^2 accuracy and simple χ^2 biases

Given this simulation setup, we generate a set of 81 (3x3x3x3) simulations forming a grid of all permutation of the following parameters: $\alpha \in [0.1, 0.15, 0.2]$, $\beta \in [2.5, 3.5, 4.5]$, $\gamma \in [0, 0.1, 0.2]$ and $\sigma_{int} \in [0.07, 0.12, 0.17]$. Each time, we draw 2,000 SNe Ia and we simulate each permutation 5 times to get, at the end 405 simulations. For each of these simulation, we fit for α, β, γ and σ_{int} for both the simple χ^2 and the total-χ^2 methods. The regular χ^2 minimisation assumes $x_{obs} = x_{true}$, while the total-χ^2 method fits for x_{true} as nuisance parameters (cf. Sect. 4.2).

We present in Fig. A.2, the results of these tests. The result are shown in the form of pulls, i.e. (fit − truth)/fit_err. If the fit is accurate, the pull distribution should be centred on zero (unbiased) with the width of 1 (correct fit error). Figure A.2 clearly illustrate that the simple χ^2 results are biased toward lower values (closer to zero) for all standardisation parameters, especially for β. This is because the typical measurement errors ($\overline{\sigma_c} = 0.03$) are non-negligible in comparison to the typical underlying colour scatter (STD(c) = 0.15). On the opposite, standardisation parameters are always accurately recovered when using the total-χ^2 used in this analysis and detailed in Section 4.1.

## Appendix A.2: Significance of the broken magnitude-stretch relation

We use our simulation tool to further test the veracity of the broken-α model. We simulate data using a single α and fit for a broken relation that split at the expected $x_1 = -0.5$, and then compute the odds to find such a strong difference in α between the low and high-stretch modes.

We thus simulate 1,000 samples of 927 SNe using a linear α, setting $\alpha = 0.17$, i.e. close to that obtain using our sample while fitting a single α, and then perform the broken-α fit. We show in Fig. A.3 the resulting distribution of the fitted α_{low} and α_{high}. They are accurately recovered in close agreement with the input $\alpha_{low} = \alpha_{high} = 0.17$. The $\alpha_{low} = 0.271 \pm 0.011$ and $\alpha_{high} = 0.083 \pm 0.009$ fitted on our data (teal point in Fig. A.3)

---

[1] skysurvey.readthedocs.io
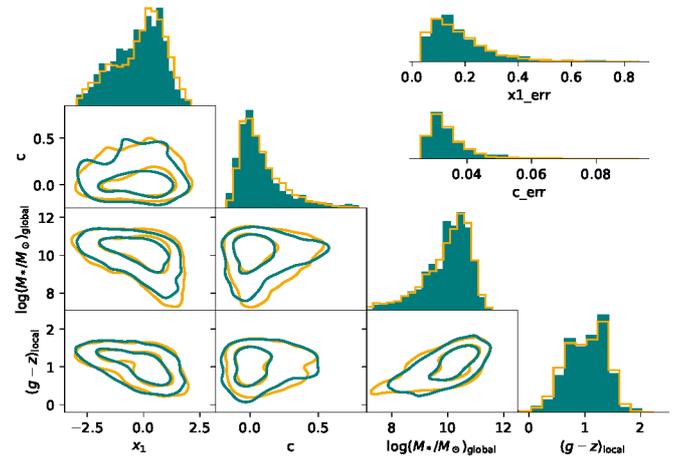
Fig. A.1. *Corner plot:* Correlation between SN Ia lightcurve stretch (x_1), colour (c), host mass and environmental local colour for the data (teal) and our realistic simulations (orange). Lines in the non-histogram panels show the contours containing 90% and 50% of the data (teal) and simulated points (orange). *Top right:* Distribution of the stretch and colour errors following the corner-plot colour code.
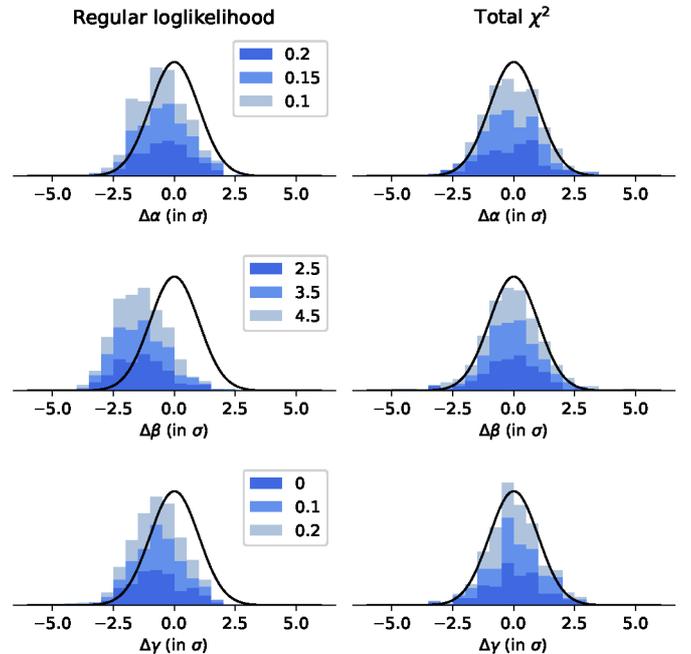

Fig. A.2. *Left:* Difference between the fitted (α, β, γ) and the input value when using a usual log-likelihood minimisation code, in σ. The stacked histograms corresponds to three different input values defined in the legend on the right. The black line is the model the histogram should follow if the results were unbiased (both values and error), which is a Gaussian with a null mean and unitary standard deviation. *Right:* Same as the left plot, but for the total-χ^2 method.

are thus confirmed to be significantly different, confirming the non-linearity of the residuals-stretch relation.

Appendix A.3: Difference between the step fitting methods

We investigated the step values obtained with the different methods described in Sect. 5, to validate that the discrepancies seen between the steps were due to variations in the fitting procedures.
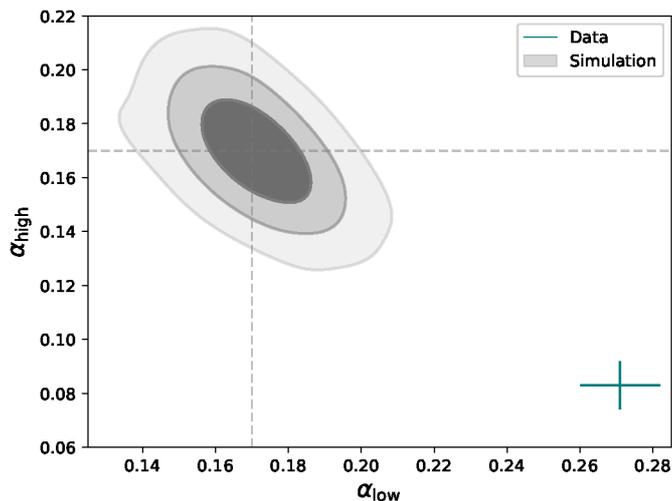
Fig. A.3. α_{low} vs α_{high} when fitting a broken-α on samples simulated with a linear α. The dashed lines represent the input values. The contours (respectively 50%, 84% and 97%) are computed with 1,000 simulations of a similar-sized sample than the one used throughout the paper, and the data point represent the values from the fit in Sect. 4.5.
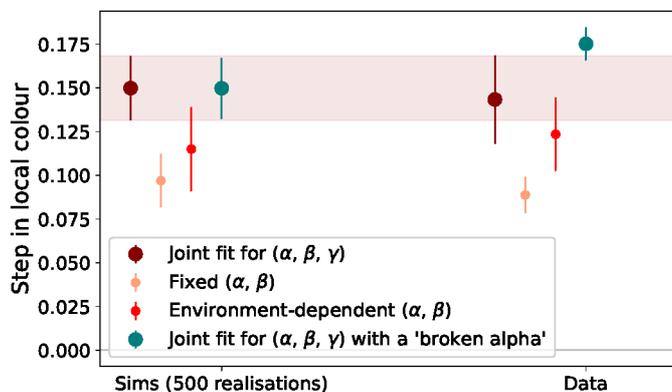
Fig. A.4. Local colour step for the four fitting procedures presented in Fig. 10, with 500 realisations of the simulations described in Sect. A.3 (left) and data (right). The points for the simulations are the means of the realisations, and the errorbars are the standard deviation.

We used the simulations presented in Sect. A, with a broken-α and a local colour step of $\gamma = 0.16$ mag. The results are presented in Fig. A.4. The order of the points seen in the data is well reproduced by the simulations. Indeed, the "Fixed (α, β)" step is the smallest of the four steps, while the other three steps are all compatible. The factor between the regular step ("Joint fit for (α, β, γ)") and the "Fixed (α, β)" step is the same in the data and in the simulations. The data exhibits a step difference of $\Delta\gamma = 0.055 \pm 0.027$ mag, while the simulations exhibits a $\Delta\gamma = 0.052 \pm 0.024$ mag difference (a 0.05σ difference between data and simulations). In the simulations, the broken-α step is of the same order as the regular linear-α step, while it is slightly higher in the data, with a reduced errorbar. However, we know the link between step and environment is more complex than the current model used in simulations, and that as environment strongly correlates with stretch, this will likely impact the broken-α step. A further modelling of an age step will be the subject of a future analysis.